

DATABASE RANGE CONSTRAINT THAT IS DYNAMICALLY LIMITED BY DATA IN THE DATABASE

BACKGROUND OF THE INVENTION

1. Technical Field

5 This invention generally relates to computer systems, and more specifically relates to database constraints.

2. Background Art

 Since the dawn of the computer age, computers have evolved and become more and more powerful. In our present day, computers have become indispensable in many
10 fields of human endeavor including engineering design, machine and process control, information storage and retrieval, and office computing. One of the primary uses of computers is for information storage and retrieval.

 Database systems have been developed that allow a computer to store a large amount of information in a way that allows a user to search for and retrieve specific
15 information in the database. For example, an insurance company may have a database that includes all of its policy holders and their current account information, including payment history, premium amount, policy number, policy type, exclusions to coverage, etc. A database system allows the insurance company to retrieve the account information for a single policy holder among the thousands and perhaps millions of policy holders in
20 its database.

Database constraints are known in the art, and are used to limit data that may be entered into a database column. For example, a range constraint for a column that includes the age of people may have a range constraint that limits the entries to integer values from 1 to 99. If someone tries to enter a negative number, or zero, or 110, the
5 database manager will detect that the data specified for the age column does not satisfy the limits specified in the range constraint for the age column. In this manner, data may be limited to expected ranges.

Known range constraints are static. Fixed numerical limits are selected (*e.g.*, 1-99 in the example above). If the numerical limits for a range constraint need to be changed
10 later on, the database administrator will have to manually change the limits as required. In many database systems, the natural evolution of the database requires manual adjustment of the fixed numerical limits for range constraints. For example, if a database table specifies a price range today, that same price range will likely need to be higher ten years from now, or even next year. Without a way for limits of range constraints to be
15 adjusted based on data in the database table, the computer industry will continue to suffer the high overhead of having a database administrator change each fixed limit for range constraints that needs to be changed.

DISCLOSURE OF INVENTION

According to the preferred embodiments, a dynamic range constraint is defined
20 that has one or more limits that are computed from data residing in the database. In a first embodiment, a dynamic range constraint is defined for a database column that includes at least one limit that is computed from data within that column. In a second embodiment, a dynamic range constraint is defined for a database column that includes at least one limit that is computed from data in a different database column. The limits may be computed

using any suitable heuristic or criteria, including statistical analysis of the data in the database. In this manner a range constraint for a column may evolve automatically as the data in the database changes.

5 The foregoing and other features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention, as illustrated in the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

10 The preferred embodiments of the present invention will hereinafter be described in conjunction with the appended drawings, where like designations denote like elements, and:

FIG. 1 is a block diagram of an apparatus in accordance with the preferred embodiments;

FIG. 2 shows a sample database table for illustrating the concepts of the prior art and the first embodiment;

15 FIG. 3 shows a sample prior art range constraint for column C3 in the table of FIG. 2;

FIG. 4 is a flow diagram of a prior art method for entering data into a column that includes a range constraint;

20 FIG. 5 shows a first dynamic range constraint in accordance with the first embodiment that has two limits defined by performing a computation over the data in the database column;

FIG. 6 shows the computed limits for the range constraint in FIG. 5 assuming a given mean for the column C3;

FIG. 7 shows the computed limits for the range constraint in FIG. 5 assuming a different mean for the column C3;

FIG. 8 shows a second dynamic range constraint in accordance with the first embodiment that has one limit defined by performing a computation over the data in the database column;

FIG. 9 shows the computed limits for the range constraint in FIG. 8 assuming a given maximum value in column C4;

FIG. 10 shows the computed limits for the range constraint in FIG. 8 assuming a different maximum value in column C4;

FIG. 11 is a flow diagram of a method in accordance with the preferred embodiments for entering data into a column that includes a dynamic range constraint of the preferred embodiments;

FIG. 12 is a sample database table for illustrating the concepts of the second embodiment;

FIG. 13 shows a dynamic range constraint in accordance with the second embodiment that has one limit defined by performing a computation over the data in a different database column;

FIG. 14 shows the computed limits for the range constraint in FIG. 13 assuming a given maximum value in column C3; and

FIG. 15 shows the computed limits for the range constraint in FIG. 13 assuming a different maximum value in column C3.

BEST MODE FOR CARRYING OUT THE INVENTION

The preferred embodiments define a dynamic range constraint that has one or more limits defined by data in the database. In a first embodiment, a dynamic range constraint for a database column has one or more limits defined by data within the

column. In a second embodiment, a dynamic range constraint for a database column has one or more limits defined by data in a different database column. The limits for the dynamic range constraints of the preferred embodiments may be computed using any suitable heuristic, including statistical analysis of the data. Detailed examples in
5 accordance with the preferred embodiments are presented below.

Referring to FIG. 1, a computer system 100 is one suitable implementation of an apparatus in accordance with the preferred embodiments of the invention. Computer system 100 is an IBM eServer iSeries computer system. However, those skilled in the art will appreciate that the mechanisms and apparatus of the present invention apply equally
10 to any computer system, regardless of whether the computer system is a complicated multi-user computing apparatus, a single user workstation, or an embedded control system. As shown in FIG. 1, computer system 100 comprises a processor 110, a main memory 120, a mass storage interface 130, a display interface 140, and a network interface 150. These system components are interconnected through the use of a system
15 bus 160. Mass storage interface 130 is used to connect mass storage devices, such as a direct access storage device 155, to computer system 100. One specific type of direct access storage device 155 is a readable and writable CD RW drive, which may store data to and read data from a CD RW 195.

Main memory 120 in accordance with the preferred embodiments contains data
20 121, an operating system 122, a database 123, a database manager 125, and database constraints 126. Data 121 represents any data that serves as input to or output from any program in computer system 100. Operating system 122 is a multitasking operating system known in the industry as OS/400; however, those skilled in the art will appreciate that the spirit and scope of the present invention is not limited to any one operating
25 system. Database 123 is any suitable database, whether currently known or developed in

the future. Database 123 preferably includes one or more database tables 124. Each database table 124 preferably includes data organized in columns and rows, with each row in the database table 124 defining an entry (or record) in the corresponding database table 124. Database manager 125 provides an interface to database 123 that allows

5 reading data from and writing data to the database table(s) 124. Database constraints 126 include known database constraints, such as primary key constraints, unique constraints, foreign key constraints, and check constraints, which include range constraints. Note, however, that database constraints 126 additionally include dynamic range constraints 127 in accordance with the preferred embodiments. The dynamic range constraints 127

10 are dynamic in the sense that they must determine one or more of their limits by analyzing data in the database. By providing limits for the dynamic range constraints 127 that are a function of data in the database, the dynamic range constraints 127 naturally and automatically evolve as the data in the database changes. For this reason, the dynamic range constraints 127 avoid making the system administrator manually change the fixed

15 limits of prior art range constraints when a value does not lie within the range defined by the fixed limits.

Computer system 100 utilizes well known virtual addressing mechanisms that allow the programs of computer system 100 to behave as if they only have access to a large, single storage entity instead of access to multiple, smaller storage entities such as

20 main memory 120 and DASD device 155. Therefore, while data 121, operating system 122, database 123, database manager 125, and database constraints 126 are shown to reside in main memory 120, those skilled in the art will recognize that these items are not necessarily all completely contained in main memory 120 at the same time. It should also be noted that the term “memory” is used herein to generically refer to the entire virtual

25 memory of computer system 100, and may include the virtual memory of other computer systems coupled to computer system 100.

Processor 110 may be constructed from one or more microprocessors and/or integrated circuits. Processor 110 executes program instructions stored in main memory 120. Main memory 120 stores programs and data that processor 110 may access. When computer system 100 starts up, processor 110 initially executes the program instructions
5 that make up operating system 122. Operating system 122 is a sophisticated program that manages the resources of computer system 100. Some of these resources are processor 110, main memory 120, mass storage interface 130, display interface 140, network interface 150, and system bus 160.

Although computer system 100 is shown to contain only a single processor and a
10 single system bus, those skilled in the art will appreciate that the present invention may be practiced using a computer system that has multiple processors and/or multiple buses. In addition, the interfaces that are used in the preferred embodiment each include separate, fully programmed microprocessors that are used to off-load compute-intensive processing from processor 110. However, those skilled in the art will appreciate that the
15 present invention applies equally to computer systems that simply use I/O adapters to perform similar functions.

Display interface 140 is used to directly connect one or more displays 165 to computer system 100. These displays 165, which may be non-intelligent (*i.e.*, dumb) terminals or fully programmable workstations, are used to allow system administrators
20 and users to communicate with computer system 100. Note, however, that while display interface 140 is provided to support communication with one or more displays 165, computer system 100 does not necessarily require a display 165, because all needed interaction with users and other processes may occur via network interface 150.

Network interface 150 is used to connect other computer systems and/or workstations (*e.g.*, 175 in FIG. 1) to computer system 100 across a network 170. The present invention applies equally no matter how computer system 100 may be connected to other computer systems and/or workstations, regardless of whether the network
5 connection 170 is made using present-day analog and/or digital techniques or via some networking mechanism of the future. In addition, many different network protocols can be used to implement a network. These protocols are specialized computer programs that allow computers to communicate across network 170. TCP/IP (Transmission Control Protocol/Internet Protocol) is an example of a suitable network protocol.

10 At this point, it is important to note that while the present invention has been and will continue to be described in the context of a fully functional computer system, those skilled in the art will appreciate that the present invention is capable of being distributed as a program product in a variety of forms, and that the present invention applies equally regardless of the particular type of computer-readable signal bearing media used to
15 actually carry out the distribution. Examples of suitable computer-readable signal bearing media include: recordable type media such as floppy disks and CD RW (*e.g.*, 195 of FIG. 1), and transmission type media such as digital and analog communications links.

Referring to FIG. 2, a sample database table T is used to illustrate the concepts of the prior art and of the first embodiment of the present invention. Table T includes four
20 columns C1-C4, and N rows, as shown. We assume for this simple example that column C1 contains a part number, column C2 contains a three letter code that identifies the supplier of the part, column C3 contains the price for the part, and column C4 contains the quantity on-hand of the part.

With table T as shown in FIG. 2, a prior art range constraint could be defined as shown in FIG. 3 that specifies a range for values in price column C3 between 1 and 25. This range constraint will prevent data outside of the range from 1-25 from being entered into the price column C3 in table T. The database manager that handles insertion of data
5 into table T will enforce the range constraint limits to keep the data in the column within the specified range.

Referring to FIG. 4, a prior art method 400 represents steps performed by a database manager once the prior art range constraint in FIG. 3 has been defined. Method 400 begins when data needs to be entered into a column that has a range constraint with
10 fixed limits (as shown in FIG. 3). First, data to be entered into the column is compared with the fixed limits of the range constraint (step 410). Thus, for the specific example in FIG. 3, step 410 determines whether the data to be entered is between 1 and 25. If the data to be entered satisfies the fixed limits of the range constraint (step 420=YES), the data is entered into the column (step 430). If the data to be entered does not satisfy the
15 fixed limits of the range constraint (step 420=NO), the data to be entered is rejected (step 440). In this manner, prior art method 400 limits the values of data that may be written to a database column by restricting data entry to data values that satisfy the fixed limits of the range constraint in FIG. 3.

The prior art range constraint, an example of which is shown in FIG. 3, includes
20 only fixed numerical limits. Nowhere does the prior art teach a limit for a range constraint that may be dynamically determined. The preferred embodiments define a dynamic range constraint that includes one or more limits that may be dynamically determined from data in the database. In a first embodiment, a dynamic range constraint is defined for a selected column with one or more limits that are computed by analyzing
25 data in the selected column. In a second embodiment, a dynamic range constraint is

defined for a selected column with one or more limits that are computed by analyzing data in a column that is different than the selected column. In both cases, the database manager analyzes data in a database to dynamically determine limits for the dynamic range constraints.

5 FIG. 5 shows one example of a dynamic range constraint within the scope of the first embodiment. The dynamic range constraint in FIG. 5 specifies that the value of column C3 must be between a lower limit that is 10% of the MEAN of column C3 and an upper limit that is 2.5 times the MEAN of column C3. We see from this example that the dynamic range constraint in FIG. 5 includes two limits that are dynamically determined
10 from data in the database. Referring to FIG. 6, if the MEAN (or average) of column C3 in table T of FIG. 2 is 10.00, the range constraint in FIG. 5 will have computed limits of 1.00 and 25.00, resulting in the same limits shown in the prior art range constraint in FIG. 3. Note, however, that the limits for the dynamic range constraint in FIG. 5 will adjust as the MEAN of column C3 changes. If the mean of column C3 is 15.00, as shown in FIG.
15 7, the computed limits for the dynamic range constraint are 1.50 and 37.50, as shown in FIG. 7. By making one or more range constraint limits a function of data in the database table, the dynamic range constraint will naturally and automatically evolve and change according to the values stored in the database. Thus, as the prices of parts increases, the ranges of allowable prices will automatically shift to reflect the increase in the mean price
20 in the column.

Referring to FIG. 8, a second example of a dynamic range constraint within the scope of the first embodiment includes a single limit that is dynamically determined from data in the database. The dynamic range constraint of FIG. 8 specifies that the value of the quantity column C4 must be between zero and twice the maximum value in column
25 C4. This type of range constraint prevents a gross error in quantity from being entered,

such as adding an extra digit or entering a negative number. If the maximum quantity in the C4 column is 150, the computed range for C4 will be from 0 to 300, as shown in FIG. 9. In similar fashion, if the maximum quantity in the C4 column is 565, the computed range for C4 will be from 0 to 1,130, as shown in FIG. 10. The first embodiment thus
5 allows defining a dynamic range constraint with one or more limits that are dynamically determined from data within the column on which the dynamic range constraint is defined.

Referring to FIG. 11, a method 1100 in accordance with the preferred
embodiments is performed by a database manager once one or more dynamic range
10 constraints (such as those shown in FIGS. 5 and 8) are defined. Method 1100 assumes that data needs to be entered into a column on which a dynamic range constraint of the present invention is defined. The first step is to compute one or more limits for the dynamic range constraint from the data in the database (step 1110). Next, the data to be entered into the column is compared with the column's computed limits for the dynamic
15 range constraint (step 1120). If the data to be entered satisfies the computed limit(s) for the dynamic range constraint (step 1130), the data is entered (step 1140). If the data to be entered does not satisfy the computed limit(s) for the dynamic range constraint (step 1130=NO), the action taken depends on whether the dynamic range constraint was specified as a warning constraint or a rejection constraint. The present invention includes
20 the concept of defining different types of dynamic range constraints. If the dynamic range constraint at issue in FIG. 11 is specified as a warning constraint (step 1150=WARN), the data is entered (step 1170), and a warning is provided (step 1180). The warning may take any suitable form, including a message to a database administrator, a return code to a program that entered the data, an e-mail to an appropriate person, etc. If the dynamic
25 range constraint at issue in FIG. 11 is specified as a rejection constraint (step 1150=REJECT), the data is rejected (step 1160). The dynamic range constraints of the

preferred embodiments thus give great flexibility in defining one or more limits that are computed from data in the database, and that may be specified as either warning constraints or rejection constraints.

A second embodiment of the present invention is presented using FIGS. 12-15. A sample database table S in FIG. 12 includes many of the columns and data shown in table T in FIG. 2, but additionally includes a cost column C4 that specifies the purchase price of the part represented by the part number from the supplier. We assume for this sample table S that the cost column C4 contains the cost paid to the supplier for the part, while the price column C3 contains the price charged to customers for the part. The difference between price and cost thus represents profit made from selling the part. The second embodiment differs from the first embodiment by allowing a dynamic range constraint to be defined on a first column that includes one or more limits that are computed from data in one or more other columns. Referring to FIG. 13, a dynamic range constraint in accordance with the second embodiment specifies a range for the cost column C4 that goes from zero to the maximum value in the price column C3. Assuming that parts are never sold below cost, the value in the cost column C4 should never exceed the maximum price in the price column C3. Referring to FIG. 14, if the maximum price in column C3 is \$18.45, then the dynamic range constraint of FIG. 13 specifies a range for column C4 from zero to \$18.45. If the maximum price in column C3 is \$78.20, the dynamic range constraint of FIG. 13 specifies a range for column C4 from zero to \$78.20. Thus, as the prices in the price column C3 rise, the maximum allowable cost for the cost column C4 goes up as well. In the specific example shown in FIGS. 12-15, the dynamic range constraint for a first column includes one or more limits that are a function of data in a second column in the same database table. Note, however, that the second embodiment expressly extends to defining a dynamic range constraint for a first column

that includes one or more limits that are a function of data in one or more other columns, which may include column(s) in a different database table.

The specific examples shown in the figures illustrate some very simple and basic computations that could be performed on data in the database to compute one or more limits for a dynamic range constraint. Of course, more complex computations could be used based on any suitable heuristic that perform more sophisticated statistical or other complex analysis on the data in the database. Performing complex computations on every database operation that has a dynamic constraint would incur some processing overhead. To minimize this effect, dynamic constraints could be configured such that their limits are computed only at set times, such as hourly, daily, monthly, etc. The present invention expressly extends to any and all methods for dynamically computing a limit for a range constraint from data in the database, regardless of the simplicity or complexity of the computation.

The preferred embodiments provide range constraints that can grow and evolve as a database grows. Let's assume, for example, that a company that sells mail order items has a price range from \$0.25 to \$89.50 for all items in its catalog. A dynamic range constraint could be defined for the price column that includes limits determined from analysis of the price data in the price column. Now let's assume this company acquires another small mail order company that sells mail order items with a price range of \$1.99 to \$149.00. As the items for the new company are added to the database, the limits for the dynamic range constraints will automatically be adjusted to reflect the newly-added data. As a result, the dynamic range constraints of the preferred embodiments provide a way to define range constraints that will minimize the likelihood that a system administrator will have to re-defined the limits for the range constraints in the future.

One skilled in the art will appreciate that many variations are possible within the scope of the present invention. Thus, while the invention has been particularly shown and described with reference to preferred embodiments thereof, it will be understood by those skilled in the art that these and other changes in form and details may be made
5 therein without departing from the spirit and scope of the invention.

What is claimed is: